4.0 MODELING AND MEASUREMENT RESEARCH NEEDS

Chapter 3 examined ATD modeling system requirements (derived from user needs) and capabilities at the level of major functional components, such as the meteorological data inputs or the transport and diffusion code. This chapter identifies ATD modeling system capability gaps and R&D options to fill them by exploring modeling and measurement science and technology in greater technical depth. The objective is to identify the most promising scientific and technical opportunities to meet the ATD modeling needs.

ATD models require knowledge of the local wind and turbulence fields at the scales of interest of the population at risk. Because these scales vary by incident and potential consequences, the domains of interest are case dependent and not usually known *a priori*. Thus, the modeler's toolbox of capabilities and the skills to use them must cover a wide range. Nearby measurements are often not available, and conditions change rapidly, especially near the ground where people live and work and are most likely to encounter airborne hazardous materials.

Advances in current ATD modeling are likely to come from improvements in meteorological model predictions and from measurements at the scales of interest. The former are closely related to better representations of atmospheric boundary layer (ABL) processes by improved parameterizations, initial conditions, boundary conditions, and representations in complex environments. As existing modeling and observing capabilities are improved, incorporating the realization that ATD processes are partly stochastic rather than entirely deterministic will enable better quantification of the uncertainties in the modeling process. The modeler must then learn how to communicate this uncertainty information to end users in ways that are relevant to the users' decisions.

Models and data must come together and complement one another. Techniques to localize and/or quantify source characteristics by fusing information from concentration sensors, ATD models, and other measurements are lacking or untested. To meet user requirements for timely modeling predictions, faster methods are needed to determine the quality of observed data, merge the acceptable data into modeling frameworks, and estimate concentrations rapidly across several scales of motion. Finally, to ensure the quality of the model estimates and provide the benchmark for improvements, the skill of the prediction and its robustness need to be assessed on a continuing basis.

Section 4.1 explains the methodology applied by the JAG to prioritize the R&D requirements and opportunities presented here and in chapter 5. Section 4.2 focuses on modeling methods to address model deficiencies and unmet needs noted in Chapter 3. Section 4.3 does the same for measurement technology, including advanced sensor systems and methods for improving meteorological data inputs. Section 4.4 focuses on capability challenges related to the *interface* between data inputs and the ATD code. Understanding (and meeting) these interface challenges requires viewing them from both the data side and the modeling side (parameterization techniques, algorithm development, etc.).

4.1 R&D Prioritization Methodology

The JAG considered the following factors and associated questions when prioritizing R&D needs and opportunities.

- **Time sensitivity.** Is there a window of opportunity for achieving results? Does other R&D depend on this work (is it a prerequisite)? Is the user need that would be met a national priority of immediate concern, or is it a longer term improvement (longer term need)? The three values used for time sensitivity are *immediate*, *near term*, or *longer term*.
- **Short-term gains**. Can the R&D results be ready for transition to operations within 2 years of initiating the R&D effort? For the research needs discussed by the JAG, short term gains were rated as either *minimal, average,* or *high*.
- **Overall level of effort (LOE).** What are the total resources that the JAG members anticipate will be required relative to other R&D needs in this plan? Specific dollar amounts or ranges (i.e., quantitative cost estimates) were not considered. Instead, this factor includes the JAG's qualitative estimate of the relative scale of labor, infrastructure, and procurement costs. Within the research needs discussed by the JAG, the LOE was designated as either *low, moderate,* or *high*.
- **Lead time.** Is this a long-lead effort; i.e., an effort that must be planned and initiated a relatively long time before an initial operational capability can be realized, or could a coordinated effort started quickly reap benefits soon? Lead times were rated as either *short* (within 2 years), *average* (more than 2 years but less than 7), or *long* (up to 10 years).
- **Ultimate potential for gain.** What is the ultimate potential for gain relative to other research needs? Because all of the R&D needs selected by the JAG for this plan were considered above average in their ultimate benefits, the three ratings used were *above average*, *high*, or *exceptional*.

Table 2 illustrates how a hypothetical R&D need might be rated using the prioritization factors.

TABLE 2. Example of Prioritization Factor Ratings for a Hypothetical R&D Need

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	minimal	moderate	long	exceptional

- This need is not an immediate national concern, but it should be addressed soon. The time sensitivity rating is therefore *near term*.
- This need is not expected to provide short-term gains because it has a long-lead time (between 7 and 10 years) before initial operating capability is likely to result. The JAG members therefore rated it *minimal* for short-term gains and *long* for lead time.

- Directed and applied research will be required to tackle this research problem, but it is not comparable to the largest efforts considered by the JAG. Therefore, its overall level of effort was rated as *moderate*.
- The ultimate potential for gain is great—among the highest of the efforts considered by the JAG. Therefore, its rating for ultimate potential gain is *exceptional.*

Given its exceptional potential for ultimate gain and the long lead required, the R&D to address this hypothetical need should be started as soon as possible. A carefully coordinated R&D plan should be developed to control cost, ensure that the potential gain is realized, and provide for ongoing evaluation of the requirement and the R&D direction over time.

As another example of the prioritization scheme, consider the kind of R&D effort that is sometimes referred to as “low-hanging fruit” because the benefits can be acquired relatively easily and quickly. An R&D effort of this kind might be rated as *high* for short-term gain, *low* for overall level of effort, and *short* for the lead time required. The “low-hanging fruit” metaphor is typically applied to something of value but not essential to have immediately or of the greatest ultimate importance. Its time sensitivity would therefore probably be *near term* or *longer term,* and its ultimate potential for gain might be *above average* or *high*.

4.2 Improving ATD Meteorological and Concentration Models

4.2.1 The Meteorological Model Components of ATD Modeling Systems

Predicting the concentration of airborne material at a given location and time after a release from a given source—the purpose of ATD modeling—cannot be isolated from predicting wind and turbulence, which is what a meteorological model does. The equations for conservation of mass (prediction of concentration given the source) of the airborne material are the same as for other scalar atmospheric variables, such as specific humidity and potential temperature. The sources and sinks (decay, chemical transformation, deposition) may vary by the material, but the movement of the material is controlled by the local wind and turbulence fields.

Because the process of estimating the wind and turbulence in the areas of interest to ATD is largely independent of the source term release event, the ATD modeling process is usually divided into a meteorological model and a concentration model. The former represents the wind and turbulence; the latter represents the relationship between source and concentration at a location given the meteorological conditions. When identifying the R&D needed to address capability gaps in ATD modeling, the capabilities of the meteorological modeling component of the system must be included.

For any given realization, an environmental prognostic model can depict only about two decades of distance scales. Therefore, as finer resolution is sought, the domain covered by a single realization must shrink. In four-dimensional atmospheric models, increasing the resolution by halving the grid sizes leads to at least an order of magnitude increase in computational burden. Distributed processing has helped reduce this burden but cannot eliminate it. Increased computational capabilities have enabled nested prognostic models to be run at lateral grid spacings of hundreds of meters. Model nesting, however, raises issues of its own, which are discussed below.

In the cascade of atmospheric energy toward smaller-scale motions, more information is required about issues at smaller scales. Instead of worrying about a few large processes, the modeler must contend with a multitude of small ones. Misrepresentation of actual processes with approximate expressions induces error and inhibits the quality of the model results. In Figure 7, moving from right to left represents the traditional top-down approach to modeling. The "bottom-up" processes of examining the flow from smaller scales, through physical modeling or simplified high-resolution numerical models (discussed in Appendix C) are represented in Figure 7 as progressing from left to right. In the range of scales from tens of meters to a few kilometers, the models do not adequately replicate atmospheric motions.

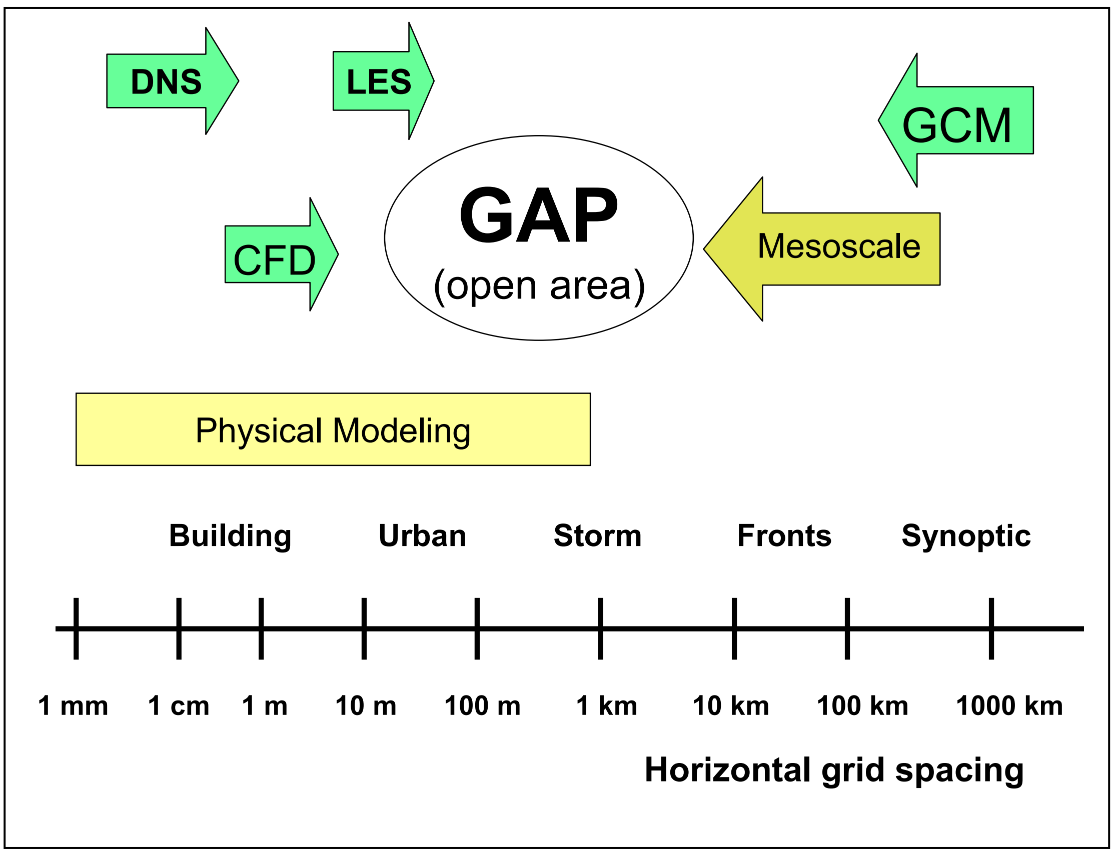

FIGURE 7. Transport and diffusion scales and model grid sizes. GCM = global climate models. Courtesy of T. T. Yamada.

4.2.2 R&D Needed to Improve ATD Modeling Components

This section identifies R&D requirements based on the preceding discussions of ATD modeling concepts and current capabilities. The rationale for each requirement follows the statement of the requirement. The rationale is followed by a prioritization assessment of the R&D to meet the requirement, using the prioritization factors and methodology introduced in section 4.1.

R&D Need: Bridge the gap from mesoscale to microscale/urban scale.

Deterministic modeling at grid scales that would allow representation of transport and diffusion phenomena characteristic of urban regions or subregions (100 m to 1 km) does not seem feasible in the near term. The accumulation of error limits the applicability of a top-down approach beginning with a mesoscale meteorological model nested in a synoptic model. The complexity of the near-surface environment requires finer and finer detail about surface features and their temporal changes. Although part of the problem may be addressed by more accurate approximations of sub-grid processes, it is difficult to develop estimation techniques that are sensitive to every nuance of a complex, poorly quantified feedback system like the urban atmosphere.

Top-down modeling is limited in its ability to represent ABL processes. An ABL process, such as three-dimensional, heterogeneous, anisotropic turbulence that scales with the boundary layer height Zi (about 1 to 2 km), cannot be represented with lateral grid scales of a few kilometers, regardless of the vertical grid spacing. In convective conditions, the lateral scales are about 1.5 times Zi. Consequently, there is a gap in capability between top-down modeling represented by mesoscale meteorological modeling, and bottom-up modeling represented by physical models and CFD models and LES. Unfortunately, the gap lies at the scales of phenomena that affect people. Considerable innovative thought is needed to bridge this modeling gap.

Useful improvements in nesting or initialization approaches include better forecasts of boundary layer height and the wind speed and direction profile as a function of grid size. The data from 915 MHz wind profilers, now available in many locations around the country, can provide the ground-truth data. These data are particularly useful for ATD modeling where ABL processes are often dominant. If the information requirements and ground truth are well defined, remote sensing could aid in providing input data to models and in model performance evaluation.

TABLE 3. Prioritization Factors for Bridging the Modeling Gap from Mesoscale to Microscale

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	average	moderate	average	Exceptional

Bridging the modeling gap is a long-term objective that will not be easily met. Success in bridging the two decades of length scales that are still poorly measured (the gap area in figure 7) depends on meeting the other R&D needs for ATD modeling identified in this section. The initial efforts to quantify the uncertainty in existing meteorological models, fine-scale models, and physical models can be used to identify areas where more extensive research and model development are needed. Longer-term progress will depend on creative initiatives producing new, testable hypotheses about boundary-layer behavior and surface layer turbulence. This longer-term effort is essential to complete the modeling of the full spectrum of atmospheric processes at work in transport and diffusion.

R&D Need: Improve characterizations of surface boundary conditions in model parameterizations and in input data sets (initial and boundary conditions).

Accurate, well-resolved data on local surface conditions are critical to credible solutions of the equations describing ATD in the ABL. As noted in sections 3.2.2 and 4.2.3, key variables of interest are related to surface energy budgets and their spatial and temporal variations. Hence, information on surface type, surface cover and condition, surface temperature, surface moisture, and other characteristics is essential. In urban environments, the data must include accurate and up-to-date depictions of the buildings, as well as the often surprising amount of greenery and its effects on surface moisture and temperature. The three-dimensional distributions of surface characteristics and effects in urban environments are important because of the strongly three-dimensional nature of the wind, turbulence, and temperature fields in these areas. Remote sensing from aircraft and satellites will probably be the best solution (see section 4.3.2); however, ground-truth data and methods to test and correct the set of data intended as the initial conditions for ATD modeling will continue to be critical for improving accuracy and quantifying uncertainties (see section 3.2.3, ATD Input Processing). Further research is required to determine how accurate this description must be and the effects of scale on model performance. For CFD and laboratory (physical) models and perhaps for the next generation of high-resolution mesoscale models, the spatial resolution of the surface-boundary conditions will probably be on the order of a few meters.

Characterization of surface-boundary conditions will require collecting data over a wide range of known conditions in order to create a reliable ensemble of realizations for many possible circumstances. This effort must overcome a strong tendency in data collection experiments to look only locally in assessing surface energy characteristics and fluxes. Surface flux measurements are time-averaged values. As the averaging time increases, larger motion scales contribute to the measured fluxes. The height of the measurement above ground is important because of the upwind influences on the measurements. In stable boundary layers, near-ground stratification may separate the surface–air interactions from the processes used to measure them and from the sensors intended to measure them, (Mahrt, et al, 2001). Businger (1989) stated the issue clearly: "…for reliable flux measurements near the surface, we need to know the height of the convective boundary layer, the entrainment at the top of the boundary layer, and the mesoscale divergence and advection. We have hopes that, with remote sensing in the boundary layer, a significant portion of the required knowledge will be obtained."

TABLE 4. Prioritization Factors for Improving Characterization of Surface Conditions and Input Data Sets

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	average	high	average	exceptional

Inclusion of local surface heterogeneity in initial conditions for ATD modeling will be a challenging objective. In the short to near term, capabilities exist to describe surface morphology at high resolution using airborne and space-based platforms. These capabilities are valuable for routinely identifying morphology changes over time. Translating these surface morphology data into representations of surface energy and momentum budgets at high resolution in day and night conditions is the more daunting challenge. Because many complex processes contribute to these budgets across a range of time scales, the large uncertainties in estimates will be hard to reduce. It will probably be necessary to continue the R&D effort over the long term with modest improvements in capability being achieved along the way.

Initial and lateral boundary conditions need to be more representative of the local environment. Progress will depend on the improved capability to characterize flows both laterally and vertically in the ABL through measurement at the scales of interest.

R&D Need: Test and refine the physical basis for sub-grid-scale parameterizations.

At each step in scale from the global to ABL modeling domains, approximations are made to account for physical processes that are too fine in scale to be resolved at the scale of the model. (Section 3.1 discusses the basics of model scales.) These approximations, called *sub-grid-scale parameterizations*, represent the unresolved processes using resolved grid values. They provide closure to the model (i.e., the model is closed when it has the same number of equations as independent variables). Because they are generalizations, sub-grid-scale parameterizations introduce sources of error. The cumulative effect is that the errors propagate (grow) as model output from one scale is passed as input to a model at a smaller scale. Well-designed models suppress the growth of accumulated errors, forcing them to dissipate; however, both theory and observation show that energy can transfer from small to larger scales. The consequences of such processes on a model realization are often suppressed or eliminated by the very techniques used to dissipate the errors from sub-grid-scale parameterization. Thus, two issues face the R&D community:

- How large are the errors from sub-grid-scale parameterization?
- How can these errors be reduced without interfering with accurate representation of real processes of energy transfer from one scale to another?

While the problems have been recognized, very few attempts have been made to resolve them. Recent basic research suggests that a bottom-up approach—representing the finer-scale processes at high resolution first and then generalizing to larger scales—allows the available data to be used effectively while improving the parameterizations that must be included for closure. Physical models provide a mechanism for measuring at very fine scales under controlled conditions so that the parameterizations can be suggested by data or physical insight and tested independently. The consequences of averaging small scales to accommodate larger grid sizes can then be evaluated. High-resolution computational models can be (and have been) similarly used to understand and improve parameterizations.

One suggested approach is to use grid-filtered equations rather than Reynolds stress models to assess the surface-layer energetics. The sub-filter-scale representations of these processes appear to have a wide range of stabilities (Sullivan et al. 2003).

A second approach uses the characteristic that atmospheric processes near the surface tend to scale with the height above ground level. One-dimensional closures connecting the first level or two of the model to the surface processes can be tested against observations, as Poulos and Burns (2003) have done. Their work showed significant scatter—suggesting unpredictability—in the Louis surface parameterization for stable boundary layers.

A third approach is to measure the atmosphere or concentrations at high resolution in all four dimensions. This will require instruments with greater capabilities than currently exist. As this cannot be done everywhere, selective experimentation will be required both routinely and in focused field campaigns. A common result of field programs, using new instrumentation, is the discovery of new phenomena and new insights into how the poorly resolved processes actually behave.

The problem of parameterizations for sub-grid or sub-filter scales was also recognized by the 11th Prospectus Development Team of the U.S. Weather Research Program in its final report on meteorological research needs for improved air quality forecasting (Dabberdt et al. 2004). Efforts undertaken as part of the U.S. Weather Research Program should be closely coordinated with R&D on meteorological models to support ATD modeling systems, since the atmospheric processes to be dealt with are the same in both areas of application.

TABLE 5. Prioritization factors for Testing and Refining the Physical Basis for Sub-Grid-scale Parameterizations

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
longer term	average	moderate	average	exceptional

Improvements in models and model components will largely depend on addressing the other R&D needs in this report. Models, both meteorological and ATD, will improve incrementally with a moderate level of effort as the recommended actions are taken to quantify uncertainty, capture existing data, establish ATD test beds, and improve measurement technologies. The potential for gain in improving decision aids for users is substantial, but it will be achieved incrementally as the state of the science is advanced through methodical R&D.

R&D Need. Characterize dispersion in complex environments.

The phrase “complex terrain” is often taken to mean mountainous or hill-valley terrain, but it can apply to any terrain that affects the wind and thermal structure of the atmosphere in ways that make concepts and predictive models based on homogeneous conditions no longer appropriate. Because the interactions of terrain features with atmospheric phenomena are really the point, a better term than “complex terrain” is “complex environment.” For example, meteorologically speaking, coastal regions near large bodies of water are complex environments because the land–water interfaces often gives rise to land–sea breeze regimes. Similarly, dry regions adjacent to well-irrigated lands will generate localized wind fields. In mountainous areas, terrain steering, wind deflection, and irregular patterns of surface heating and vegetation will give rise to very complicated flow patterns. This is especially true at night when cold-air drainage into mountain valleys can produce significant jets of air moving above neighboring flatlands or basins at a time of day when the surface boundary-layer characteristics are particularly difficult to predict. In general, the common feature of all complex environments is the poorly understood impact of heterogeneous surface cover and surface energy budgets on local wind, turbulence, temperature, and moisture fields.

Urban areas are a particular focus of this report because CBRN events can be expected to occur primarily in urban settings. The enhanced roughness and changes in thermal characteristics of the urban landscape alter the meteorological fields over the city. At night, the thermal characteristics often lead to a heat island over the city, which can generate its own flow fields under light synoptic wind conditions. Around the central business district, where tall and large structures are usually clustered, one can expect to find flow deflections, flow channeling along street canyons, preferred sites for recirculating flows, and other organized flow patterns, such as wake vortices, strong vertical mixing, and greatly enhanced turbulence. To improve our understanding of complex urban environments, efforts are required to compile adequate building morphology data and to conduct useful field measurements to aid in validating and better parameterizing ATD modeling systems for urban environments. Thermal remote-sensing data can aid in documenting urban surface energy budgets by providing accurate observations of the thermal variation of urban landscapes at spatial resolutions from several meters to a kilometer.

Several urban ATD field studies have been conducted in the past 20 years, both in the United States and abroad. Despite their obvious importance, the number of these studies is small because of the expense and logistical problems in conducting field studies in actual urban areas. Our understanding of flow and dispersion in urban environments containing complex building clusters and street canyons is still inadequate. A major constraint has been the cost of making a sufficient number of in situ measurements at many locations over an extended period, especially measurements of the vertical profiles of key variables such as wind, turbulence, and temperature. Improvements in remote-sensing systems may lessen this constraint considerably while providing more representative data. Future urban studies should make extensive use of remote-sensing systems, such as radars, light detection and ranging (lidar), and sound detection and ranging (sodar), to provide both meteorological and tracer concentration data. Aircraft and satellite observations may become especially useful in understanding the processes controlling ATD over, around, and through cities. This understanding is crucial to improving predictive models. For an optimal fit between a remote-sensing system and an ATD model, both the model criteria and the phenomena to be observed must be thoroughly defined.

In a sense, the stable boundary layer (SBL) provides another type of complex environment. When an SBL is present, deep convective plumes do not exist to dilute hazardous materials released near the surface by mixing with higher-elevation air. Because of the SBL's strong surface stratification and weak or intermittent turbulence, the material remains concentrated near the surface for extended periods. SBLs occur almost nightly. Hazardous releases occurring at night (e.g., the fertilizer plant accident in Bhopal, India, in 1984) can expose large populations to concentrated doses of the hazard. Yet, SBL behavior is difficult to observe, generalize, and simulate because the weak, stratified turbulence of an SBL can be induced or maintained by a variety of processes, such as breaking gravity waves, Kelvin-Helmholtz instability, density current, or low-level jets (Banta et al. 2002). Recent high-quality field campaigns, such as CASES-99 and VTMX, have explored SBLs. Numerical studies at several scales, such as the GABLS LES study (Beare et al. unpublished) and the GABLs single-column model study (Cuxart et al. unpublished), have attempted to improve SBL parameterizations by defining intermodel variability and uncertainty due to numerical parameterizations.

SBLs are currently poorly parameterized in mesoscale models. The commonly used parameterizations (e.g., Louis 1979) frequently fail in two ways (Poulos and Burns 2003). First, the parameterizations often lead the model to predict too-rapid cooling of the surface, which suppresses turbulent mixing inappropriately. Second, they often maintain well-mixed layers that last too long and are too deep. In short, these parameterizations do not account for the intermittent sources of turbulent mixing mentioned above. Although the research community in boundary-layer meteorology is addressing the problem, substantial additional work is needed.

TABLE 6. Prioritization Factors for Characterizing Dispersion in Complex Environments

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	average	high	average	high

Initial efforts to quantify uncertainty in ATD modeling of complex environments will produce some near-term gains. Some efforts are already in progress, using existing tracer data from field experiments in cities and near missile ranges. One near-term activity should be to examine the existing archives of complex-terrain studies; however, field trials may not be sufficient to develop meaningful ensembles of realizations. Physical models of urban areas should be used to study uncertainty issues. Development of models for nocturnal transport and diffusion in cities can proceed from recent limited tests. For other environs, substantial R&D effort will be needed to understand and parameterize the nocturnal boundary layer in open, hilly, and mountainous terrain, as well as along coastlines. Significant advances in measurement technology may be needed to develop appropriate databases of meteorological and tracer data.

R&D Need. Develop methods and technologies for improving ensemble construction and interpretation.

Given the errors and uncertainties in initial or boundary conditions and in the model parts (e.g., numerical core, sub-grid-scale parameterizations), each run of a model produces a single realization in the ensemble of possible analyses (for a diagnostic model) or forecasts (for a prognostic model). As explained in section 3.2.2, an ensemble of possible realizations can be created by two basic approaches. The same model can be used to produce multiple realizations by perturbing initial conditions, using variable parameterization schemes, combining these two approaches, or using variable grid resolutions. A second approach is to use different models to produce the multiple realizations for the ensemble.

Sometimes the consensus (least different) realization in an ensemble is taken as the most likely solution. For weather forecasting, Fritsch et. al. (2000) found that the consensus realization gives better skill scores in large-scale flows than do other approaches to selecting the best weather forecast in an ensemble produced from multiple models. Although the consensus approach is now widely used and accepted for weather forecasting, its applicability to ATD modeling and the phenomena that become important at finer scales has not been established. In large-scale and strongly forced flows like hurricanes or severe storms, the ensemble approach gave better skill scores by various measures; however, the suitability of various ensemble approaches has not been established for weakly forced or ABL flows. The consensus technique has only rarely been used to examine ABL flows over the variety of surface morphology, diurnal conditions, and climatic regimes routinely encountered for ATD modeling. The JAG could find only one instance of the use of consensus selection in an advanced ATD model (HPAC) to estimate concentrations. Even for this one ATD model, there are limited comparisons of predictions using the consensus realization with observations. The scientific bounds and applicability of this technique within ATD modeling need careful experimental and theoretical study.

The complexity of near-surface flows and mixing suggests the variability may be large, and increased skill may be hard to demonstrate. Integration of local, near-surface measurements into larger-scale flow regimes has not been effective; the model and the observations often disagree about the distribution of mass (the pressure field) and momentum (wind and turbulence fields) because the scales of distributions contained in the model are much larger than the scales of the local observations. As local data become available at higher resolutions, the modeling approach of using large-scale forcing to drive local-scale models will have to bridge the scale gap. With sufficient local data, shorter-term, locally based predictions of wind and turbulence fields—and thus, predictions of concentration fields—will become feasible and reliable.

The top-down and bottom–up approaches to assessing data representation and assimilating data into models are important across the spectrum of motions. The top-down approach is driven by nesting of models to approach smaller scales. The bottom-up approach is driven by local measurements to analyze, diagnose, and predict present and future local conditions. At present, with a few exceptions, local data sources are too sparse. Remote measurement capabilities appear to be the best R&D pathway to meet this need to characterize the local wind fields for purposes of improved transport and diffusion prediction. Ensemble assimilation techniques must be developed for the local data utilization.

As noted in section 3.2.2, substantial R&D on ensemble methods is also needed on the following topics:

- The optimal number and types of ensemble methods to produce statistically significant improvements in results;
- Advanced techniques for creating the individual realizations in the ensemble;
- Development of techniques to link ensemble mesoscale meteorological prediction systems with ensemble ATD predictions and evaluate the overall uncertainty of the probabilistic results; and
- Communication of probabilistic information from ensemble techniques to users through user-tailored decision aids.

TABLE 7. Prioritization factors for Developing Methods and Technologies to Improve Ensemble Construction and Interpretation

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	**minimal**	**high**	**short**	**exceptional**

The ability to quantify the uncertainty in ATD predictions will provide invaluable guidance on where R&D resources (funding and talent) should be invested to optimize the return. Because of its tremendous potential for gain in improving the quality of ATD prediction products if used to guide other R&D investments, the need for improved ensemble construction and interpretation is immediate. The overall level of effort, however, will be substantial. Many aspects, such as characterizing the ABL or quantifying uncertainty in data-sparse environments, are likely to require an extended period of R&D. Nevertheless, even incremental improvements in our capability to quantify uncertainty will produce moderate to substantial product improvements for users of consequence assessment systems.

R&D Need. Develop and test techniques to better estimate wet and dry deposition and chemical interactions.

As explained in section 3.2.5, removal or transformation processes and resuspension of previously deposited material are important factors for predicting concentrations downwind and are critical for assessing related impacts on the terrestrial and aquatic environments. Improvements in the parameterizations of the model physics for these processes, together with better empirical coefficients, are needed to improve ATD model predictions.

Although both dry and wet deposition of airborne materials has been studied for decades, the airborne hazards of interest have generally been either radioactive products of nuclear testing or common air pollutants, such as sulfur, nitrogen, and mercury compounds. Removal or resuspension of other materials, especially small particles, is still difficult to predict. Wet deposition, whether by rainout or washout, is far more efficient than dry deposition, but it occurs only during periods of precipitation. Until more is known about these processes, especially as they occur in urban areas, improving modeling approaches to account for them will remain difficult. However, the importance is high because hazardous substances deposited in urban areas will find their way into run-off, wastewater, and aquatic ecosystems, with health and ecological effects that might be severe.

Focused studies of the both wet and dry deposition processes are needed. Additional process-level understanding is important for developing successful simulations. In the case of dry deposition, the necessary measurements and modeling will be particularly difficult in urban areas because of the heterogeneity and general complexity of the surface. For wet deposition, it is known that urban areas modify the precipitation regimes downwind, but it is not yet known whether this translates into more rapid scavenging of hazardous materials contained in the air. Carefully designed experimental programs will be required to address these questions. In addition to improving the representations of the contributing processes in models, accurate prediction of deposition, transformation, and resuspension effects requires the capability to merge real-time information (e.g., radar-derived values for rainfall rates and their spatial and temporal distributions) into the models. Prediction of exactly where precipitation will occur and at what rate requires characterization of many stochastic systems. At urban/local scales, predictions of where it will rain and at what rate are therefore likely to be best described in probabilistic terms.

TABLE 8. Prioritization Factors for Develop and Test Techniques to Better Estimate Wet and Dry Deposition

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	average	moderate	Average	high

Deposition from the atmosphere constitutes the linkage between atmospheric concentrations and ecological (and other environmental) effects. There is accordingly a long history of research on both wet and dry deposition of various substances, usually for periods of an hour or more. However, substantial uncertainty exists in current representations. Wet deposition models require significantly improved cloud and precipitation models for time, location, and intensity, with due allowance for the sometimes dominating role of processes that cannot yet be addressed deterministically. Dry deposition formulation will necessarily need to take similarly stochastic factors into account, in this case related to the heterogeneity of the surface. The contributing processes have been studied principally for contexts other than atmospheric turbulence and diffusion. The results will certainly depend on the substance being deposited.

This may be a difficult research effort, but the chances of success have improved with recent development of new measurement technologies. Use of remote-sensing methods (especially radar) and modern methods of chemical analysis could advance our understanding of wet deposition and its temporal and spatial characteristics. Likewise, new methods for measuring dry deposition rates are now ready for exploitation, particularly modifications of well-known and long-proven eddy correlation techniques.

4.2.3 Approaches for Model Improvement

Physical Model Simulations of Transport and Diffusion

Mathematical models of transport and diffusion must make substantial approximations for some of the fundamental fluid-dynamical processes involved, particularly those processes unresolved by the model. Numerical modeling can be supplemented and improved by using physical modeling (e.g., wind tunnels, water channels, or water tanks) to simulate the atmospheric conditions of interest. In these laboratory simulations, the primary variables can be controlled, and the time and expense are greatly reduced compared with full-scale field studies.

A physical model must duplicate certain nondimensional parameters if it is to provide a realistic simulation of ABL processes (EPA 1981; Snyder 1972). Unfortunately, not all of these dimensionless quantities can be matched simultaneously to their full-scale (atmospheric) values. Research has been conducted to provide advice on which quantities are most important for simulating ATD of neutral and positive buoyancy gases (EPA 1981) and of dense gases (Meroney 1986).

Laboratory experiments (wind tunnel, water channel, or water tank) have been used to investigate a number of ATD problems, including transport around individual buildings and industrial structures, through clusters of buildings, and in street canyons. Some of these have investigated the dependence of concentration fluctuations on the initial size of the source. Water channels have been used to investigate stable and neutral boundary layer transport around isolated hills. Water-tank experiments have been used to investigate convective diffusion and plume rise from explosions. Results from such laboratory simulations are the basis for many of the model parameterizations in current use for these situations.

Although laboratory simulations cannot fully replicate every characteristic of the full-scale condition of interest, they provide a cost-effective solution for exploratory research, confirmation of theoretical solutions, and construction of operational model parameterizations and estimation methods. Laboratory simulations are especially effective for investigating and characterizing the stochastic effects of ATD inherent in microscale flows around obstacles.

High-Resolution Modeling of Turbulent Flow

The fluid-modeling community has many years of experience in modeling turbulent flow regimes in a variety of circumstances. Special attention has been directed to turbulence near the interface of a fluid flowing past a fixed or movable surface, which is conceptually applicable to turbulence in atmospheric flows.

Direct Numerical Simulation (DNS). For air flows less turbulent than the atmosphere, DNS of carefully described, idealized atmospheric boundary flows can resolve turbulent eddies down to the molecular scale. Consequently, approximations are not required for very small eddies, but these conditions are applicable to only a limited number of ATD scenarios. Nevertheless, for turbulent flows near the ground, DNS has demonstrated important properties of surface-layer flows. The DNS approach can be used to help quantify processes and improve parameterizations for larger spatial scales.

Computational Fluid Dynamics (CFD). As noted in section 3.2.4, CFD codes, adapted from aerospace applications simulate mechanical turbulence in atmospheric flows around obstacles, particularly in urban settings. CFD grid spacings are several meters in the horizontal and vertical dimensions, so some representation of smaller scale processes is required. Inflow boundary conditions are often fixed in time, and larger scale motions are not usually included. In some cases, simple time variations can be imposed. In urban studies, CFD codes sometimes are embedded within prognostic mesoscale meteorological models, which provide time-varying boundary conditions for ATD simulations. In many cases, CFD codes do not account for the local sources or sinks of heat or their time variations.

The computationally intensive CFD approach can be used to study features of the complex wind fields in urban environments such as those found in the MUST, URBAN 2000, and Joint Urban 2003 field studies. An important feature of numerical simulations is that the external conditions can be controlled for many model runs, each of which has a slightly different initialization. The resulting ensemble of realizations can be used to obtain quantitative estimates of uncertainty in predicted concentration fields. Because CFD models have limited volume domains, their boundary conditions are often assumed rather than being calculated from larger scale simulations.

Large Eddy Simulation (LES). Historically, mesoscale meteorological models have employed horizontal mesh sizes that were much greater than the depth of the ABL. The entire turbulent energy spectrum of the ABL was therefore well below the resolution of mesoscale models, and turbulence parameterization methods were needed. By contrast, an LES has a horizontal grid of 10 to 50 m and can resolve the larger, more energetic, turbulent eddies in the ABL. However, even an LES is unable to resolve the finer scales of turbulence, so a sub-grid-scale parameterization is needed to account for energy exchange between the resolved grid and the unresolved grid. Recent field experiments (Sullivan et al. 2003) and modeling studies (Chow and Street 2002; Juneja and Brasseur 1999) have suggested several sub-grid-scale turbulence parameterizations that significantly improve the popular closures.

The LES has become a common tool for investigating ATD because the statistical properties of LES results show many similarities to those of atmospheric turbulence, especially for unstable and neutral stability conditions. However, because the physics of some processes in the stable ABL are not well understood, LES is still being refined for stable conditions.

Lagrangian particle models using resolved LES wind fields are often used to characterize ATD of material from various sources. Substantially different realizations of concentrations in plumes from point sources at the same or different locations within the volume are commonly calculated. The statistically similar behavior of plume characteristics (mean and variability) as a function of stability and release height has been demonstrated (Weil 2004). As with physical models or CFD, the ability to control conditions provides the opportunity to compile ensembles of realizations for uncertainty estimation.

Although the LES approach is commonly used, it is limited because the initial and boundary conditions are usually assumed rather than based on observational or modeled inputs. Since many LESs use cyclical boundary conditions, lateral motions typical of mesoscale phenomena cannot be included.

> R&D Need: Continue the development and use of physical modeling capabilities and high-resolution computational models (DNS, LES, and CFD) to simulate transport and diffusion in boundary layer and complex flow regimes and to assess components of uncertainty of concentrations and meteorological factors

These modeling approaches are the foundation of small-scale modeling and attempts to link across the mesoscale–microscale modeling gap. They have many features that are in need of R&D efforts. The models provide a capability to specify the atmospheric state and develop ensembles of realizations. With concentration estimates, the approximate bounds to the inherent uncertainty of ensemble conditions can be assessed. We can gain significant insight and knowledge of the consequences of averaging concentrations, fluxes, and turbulence. Better parameterizations of scale-dependent processes can be developed. Quantified fields of concentrations, winds, and turbulence can be analyzed. These models will be used to test and evaluate consequences of high-resolution surface characterizations and boundary conditions. The models will be used to test new closure equations. Furthermore CFD- type modeling will continue to be used in complex geometries, so R&D on including energy budgets in the simulations is a crucial issue.

TABLE 9. Prioritization Factors for Development and Use of Physical and High-Resolution Computational Models

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	Average	moderate	average	exceptional

A significant effort in small-scale ATD modeling (physical and computational) already exists. Physical models can provide an “observational” database for assessment of ATD model performance in complex conditions. Appropriate fields from high-resolution computational models are available or are reasonably easily regenerated. Analyses of existing data sets, such as those derived from Joint Urban 2003, should begin immediately to assess uncertainty issues as a function of scale. Substantial progress should come rather quickly where data exist. The effort must be sustained, as other R&D efforts progress, to help confirm new approaches and close the modeling gap.

4.3 Improving Measurement Technologies

Measurement of atmospheric properties and processes at and below the scales of interest is essential to improvements in ATD modeling. The primary area of concern of ATD is near the ground—in the surface layers of the atmosphere—where the hazard comes in contact with the ecosystem and its effects are felt. Meteorologically, the surface layer is connected to the large-scale flows through the ABL and is the most variable portion of the atmosphere.

The ABL is unique in that it results from the interaction of the small-scale effects of surface properties with large-scale flow fields. Furthermore, the ABL responds to diurnal solar heating and radiational cooling processes, providing a three-dimensional turbulence structure whose height during the day is on the order of the boundary layer height and during the night is a sharply stratified two-dimensional turbulence with little vertical mixing. Measurement systems need to account for the wide variety of conditions and scaling lengths that arise even in open areas.

Within urban areas, measurements become more difficult to make and then to understand because of the scales of buildings and the variations in vegetation and surface conditions. The increase in degrees of freedom challenges assumptions about relationships between measured quantities and can invalidate modeling assumptions. As the information in Table 1 (chapter 3) suggests, the ABL and the surface layer are poorly and sparsely sampled for ATD uses, both for tracer material (for the evaluation of ATD model performance) and for characterizing the environment (wind, temperature, humidity, and turbulence).

The standard meteorological measurement methods for weather forecasting have not focused on observing the *entities* in atmospheric phenomena, such as eddies at different scales that cause the short-term fluctuations in airborne hazard concentration at a given point. Scanning lidar systems can now identify eddies and other fine-scale phenomena rather than simply measuring their effects on state variables at particular points and times. These technological advances in measurement open up entire new strategies for observing the atmospheric processes and phenomena that cause the variability in local air movements and, therefore, in hazard concentrations. These new observations will be able to feed the parameterizations in improved mesoscale models. They also may be able to provide the initial or boundary conditions for much more sophisticated and realistic representations of microscale phenomena in ATD codes.

To quantify the uncertainties in ATD predictions, measurements of the distribution for a meteorological input parameter are necessary--not just a point observation of the parameter. For example, it would be preferable to have an observation-based estimate of the standard deviation in the wind direction during an increment of time, rather than a single point value for the wind direction. Existing quality standards for meteorological observations used by NOAA/NWS, for example, aim for accurate point estimates and are not adequate for capturing this information about the temporal distribution of the parameter.

R&D Need: Improve tracer materials and measurement technology.

Field experiments are the preferred means for testing and improving transport and diffusion models. Controlled releases of inert gases or aerosols simulate the transport of active agents through the atmosphere. Measurements of concentrations of this tracer material over a given period at various locations provide “ground truth” for the concentration and its time history, defining the plume dimensions and the distribution of mass within the plume. Measurements of wind and turbulence properties and other atmospheric variables in the study area are taken at the same time.

Previous dispersion experiments are useful mainly for studying the gross behavior of the effluent plume, as they provide reasonably good spatial location of the plume's path and footprint. The experiments were well suited for evaluating concentration prediction models that produced results with time and space scales similar to those of the sampling network. Tracer technology, however, has not kept pace with turbulence measurement technology. Improvements in tracer technology—both the tracer material and its measurement—are essential for assessing progress in ATD modeling.

Tracer materials are an important part of field experiments. Their release characterizes the source terms of interest to ATD modeling. The measurement of tracer concentration as a function of distance and time from release defines the impact of ATD on a released material and thus the potential hazard zones. Selection of the tracer technique for an experiment can be a tedious task.

Tracers must be *inert*—minimizing the potential health or environmental hazards. Aerosol tracers need to replicate the aerosols of interest. The tracer aerosol should closely replicate the hazard aerosol in size distribution, dielectric properties (for remote sensing), and affinity for moisture (since aerosols swell at unsaturated humidities). For gas or aerosol tracers, low cost per unit mass is desirable. Detectability at highly diluted concentrations (10^{-9} dilution) is also highly desirable.

Instruments for detecting tracer concentration must have a *rapid sampling rat*e, as variations in concentration for intervals shorter than the instrument response cannot otherwise be measured directly. If the response rate is not adequate, coarser parameterizations of fluxes and surface deposition are required, intermittency of concentrations cannot be determined, and short-period events of high concentrations cannot be identified. Instruments also need a large dynamic range of measurement to capture high and low concentration events as they occur and with equal precision. The devices should provide the measurement for real-time analysis. Low cost per sensor is needed to allow for a large array of sensors. Concentration detection by remote-sensing techniques is highly desirable for future studies because point measurements can never provide sufficient coverage near the ground and aloft.

Few tracer system components—material or sensor—satisfy all these requirements. Trade-offs (except for safety) are usually required. At present, sulfur hexafluoride is the gaseous tracer of choice for urban and short-range studies, as in Joint Urban 2003, Urban 2000, and Pentagon Shield. Rapid-response sulfur hexafluoride detectors are expensive and not readily available. Consequently, the measurement campaigns for these studies were highly labor intensive and logistically limited. Some long-range studies have used perfluorocarbons (e.g., the ANATEX experiments), but long integration times (about 12 hours) were required to accrue enough material for analysis (Draxler 1991; Draxler et al. 1991).

Assessing and guiding future improvements in ATD modeling will depend on a concentrated effort to develop tracer materials and measurement technologies that meet these requirements. Adequate tracer studies must become routine so that ATD model performance can be regularly evaluated. Adequate studies are also needed to evaluate the sources of uncertainty in both measurements and modeling.

TABLE 10. Prioritization Factors for Improving Tracer Materials and Measurement Technology

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	**High**	**moderate**	**short**	**exceptional**

A concentrated effort to develop needed tracer technologies should begin immediately. It is impossible to routinely quantify uncertainty in ATD models in the field or in test beds without controlled tracer data. This effort is critical for the overall research objectives but may not be easily achieved because progress on development of fast-response sensors has been slow and remote-sensing capabilities will depend on the candidate tracer material. Having this capability is also essential to learning how best to communicate uncertainty to users.

R&D Need: Improve boundary-layer atmospheric measurement capability.

Atmospheric quantities other than hazard concentration, such as temperature, water vapor, and trace gases are also affected by atmospheric fluctuations. Short-term changes in wind speed and direction, which strongly affect the transport and diffusion of locally emitted contaminants, are of particular interest in the case of hazardous or toxic materials for which even short-term exposures to a threshold concentration may seriously affect human health and safety.

Fluctuations in wind speed, wind direction, and atmospheric temperature are probably among the easiest data to obtain. Appropriate instrumentation has been available for decades and has been continuously improved. With the recent decrease in cost of three-dimensional sonic anemometers (which offer the advantages of fast response, low threshold, and no moving parts), it is now possible to establish continuous, around-the-clock measurement programs for wind and temperature fluctuations at most locations of interest. The accuracy of these instruments remains reasonably good for sampling rates up to 20 Hz. Data can be collected easily using inexpensive small computers. Data transfer and centralized data collection and storage are probably the main remaining technical problems because of the amount of data that can be collected from even a modest-sized wind-sensor network in a relatively short time.

Even with an ability to collect large amounts of data, data collection at discrete points may cause problems of data quality and data representation (see section 3.2.3) because of the spatial and temporal variability of the wind and its fluctuations. In principle, this issue could be handled by an instrument array that is dense both horizontally and vertically. While this solution is not feasible for routine instrument networks, it is feasible for specialized networks and field studies. The correlation length is a measure of the consistency of observations taken at one location with those at another location using common averaging times (e.g., 1 minute instead of 15 minutes). Dense networks provide an opportunity to measure correlation length and assess the density of measurements needed to characterize a location and its surroundings (urban or rural). Joint Urban 2003 is a good example of a field experiment that employed a dense wind and turbulence network.

Remote-sensing techniques offer the potential for technology solutions to ABL measurement needs. The major advantage of a remote-sensing method is that a line or volume of the atmosphere can be sampled at a known, designed (often rapid) rate without the sensor needing to be in direct physical contact with the spatial points or volumes being sampled. Remote-sensing systems produce observations either by active or passive means. In an active system, the system transmits a signal and records the direct or indirect interaction of that signal with environmental conditions of interest. In a weather radar system, for example, the reflections of a transmitted radio wave from a scatterer are captured. Passive sensors typically rely on the thermal properties of the ground or the atmosphere, without an emitted signal. A typical passive observing system detects the infrared energy naturally radiated from the environmental condition of interest. Both active and passive approaches to remote sensing typically require additional extensive processing of the received signal to obtain the desired information about the environmental condition of interest.

Active systems such as radar, lidar, and sodar have been used to measure winds, temperature, and precipitation. Ground-based radar and sodar have been available to ATD modeling for many years to measure wind profiles, but they have not been networked for operational use. Implementation of clear-air radars, like the FAA Terminal Doppler Weather Radar system, provides wind fields at kilometer increments near major population centers. Efforts are underway to make similar use of the WSR-88D weather radar network. Recent availability of eye-safe Doppler lidar systems has permitted resolution of wind fields of urban and rural domains at resolutions measured in tens of meters. These systems generate volumetric data in tens of seconds. Improved and expanded lidar capabilities are being studied actively for both research and operational use. Airborne and satellite-based radar and lidar capabilities are becoming common. For directed-research programs, moisture and ozone are now commonly measured using remote-sensing systems on aircraft or satellites. The ATD R&D community must keep abreast of these developments as they relate to mesoscale through microscale applications.

Satellite-based or airborne passive remote sensors can provide land cover information across a range of temporal and spatial scales. Passive remote sensing is well suited to gathering current data on local surface radiances, which can be interpreted into information about surface conditions. Thermal remote-sensing techniques can help document the surface energy budget by providing accurate observations of thermal changes across the landscape at spatial resolutions from several meters to a kilometer. Multispectral and hyperspectral imaging in fine bands within the same view may become a means of quantifying subtle but significant differences in surface conditions. Research must continue on methods of translating the sensed radiation signal into information on surface properties on a timely, reliable, and comprehensive basis.

Remote-sensing systems that use sophisticated scanning techniques, such as push broom or framing techniques, have the potential to probe larger volumes of the atmosphere above a region than do local sensors; however, these systems have their limitations. Their view of the atmospheric volume can be degraded by precipitation or blocked by obstacles or clouds. They can be affected by unwanted reflections (clutter) or by scattering in the sensing medium. They may also have relatively coarse spatial resolution. Some systems (e.g., sodars) produce signals that can affect humans in the immediate area and, as a result, may be difficult to use in populated areas. Some remote-sensing systems, especially research-grade systems, are not well suited to autonomous operation. They require frequent or even continuous attention from skilled specialists. This requirement makes them expensive to operate.

Remote-sensing systems generally have a high sampling rate, which means that very large quantities of raw data can be acquired in a short time. This rapid data acquisition poses potential storage and transfer problems. Interpreting the data is often not straightforward and may require considerable expertise and experience. For example, correct interpretation of the data is often complicated by the noise of turbulence effects, which generally must be removed by time-averaging to reveal the mean patterns. Ground-truth data over the sensing volume are needed to ensure that the remotely measured variables and derived parameters agree with conventional data sources.

Finally, research is needed to improve the fundamental understanding of how models can incorporate a variety of remotely sensed data.

TABLE 11. Prioritization Factors for Improving Boundary-Layer Measurement Technology

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	high	high	short	exceptional

To realize short-term gains, further development of existing boundary-layer measurement capabilities should be accelerated and on-the-shelf improvements should be incorporated into existing measurement systems. New, low-cost instrumentation developments for measuring important boundary-layer variables and possibly fluxes, should be started by exploiting existing R&D mechanisms such as the Small Business Innovative Research (SBIR) and Small Business Technology Transfer (STTR) programs or agency-specific instrument development programs. New instrument development is often time consuming and expensive, so options to expedite the process should be explored through laboratories, academia, and industry. The measurement improvements will enable other R&D needs to be met.

4.4 Model–Data Interface Challenges

In the end-to-end functional analysis of ATD modeling systems provided in section 3.2, there are a number of capability gaps that occur at the interface where data come into the modeling system. Such capability gaps could be viewed as either an input data problem (too few data, questionable data, not the right data, etc.) or a modeling problem (model representations not powerful enough to predict from the data given, etc.). Innovative approaches to filling them often amount to some combination of advanced input processing, as defined in section 3.2.3, and model improvements. Two recurring themes in these interface challenges are determining the *impact of input data uncertainties* on the uncertainty in modeling system predictions and issues of data quality and data representation, as defined in section 3.2.3. The JAG identified R&D needs for three areas of these model–data interface challenges: sensor fusion, data representation and data assimilation, and model performance evaluation.

4.4.1 Source Characterization by Sensor Fusion

A comprehensive assessment of source characterization technology—either current capabilities or R&D needs to meet application requirements—is beyond the scope of this report and beyond the charge of the JAG. However, the use of sensor fusion, as defined in section 3.2.1, to back-calculate to the estimated source term location, emission rate, and release duration is within the scope of this report.

In certain hazard release scenarios, the first indication of a release will be alarms triggered by specific sensors at varying distances and directions from the exact location of release. The identification of the location and characteristics (emission rate, duration) of the source from these alarm data are often important objectives in response management. By combining data from networked sensors with predictive models, more can be learned about a release event than could be obtained from any individual sensor or predictive model alone. Sensor fusion modeling systems are being developed to backtrack from these initial sensor data to estimate the most probable source term location and emission characteristics.

The process of integrating sensor data with predictive models is intended to result in adjusted predictions that better describe the release event than would model calculations made without the sensor data. The predictive models produce calculations based on previously available information about the source term release. Several predictive models may be run separately to produce a single prediction, or they can be used to create an ensemble of predictions. Some predictive models can estimate aspects of the uncertainty in their description of the event. Although the mathematical problem may not allow a single, definitive solution, even the capability to narrow the possible range of locations or characteristics can be valuable to response decisions.

To explore the limits in characterizing the source term by coupling ATD modeling with monitoring data, Hanna, Chang, and Strimaitis (1990) analyzed data from the Project Prairie Grass diffusion experiment using Gaussian plume models. Two of the three models had been tuned to the release situation. The authors concluded that the source term emission rate could be estimated to no better than a factor of two, even with an advanced Gaussian plume model, a point source in ideal circumstances, and a near-surface release having a known release height and a constant emission rate. This limit on emission rate estimation may represent an ultimate uncertainty that cannot be reduced, but further investigation is warranted.

R&D Need: Improve and evaluate sensor fusion techniques.

The R&D areas for improving sensor fusion techniques include the following:

- Rapid interpretation of data streams from multiple sensors;
- Increased detection confidence with reduced system-level false alarms;
- Improved situational awareness for CBRN events;
- Estimates of the most probable source terms; and
- Refined model predictions of downwind hazards.

Sensor fusion methodologies with the potential to provide these improvements include inverse dispersion modeling, Bayesian statistical methods, adjoint methods, artificial intelligence, neural networks, fuzzy logic, and others. Sensor fusion techniques should be developed that can use data from a wide range of detector types, as well as other relevant non-sensor data, such as intelligence and medical information. Existing mathematical and statistical concepts should be evaluated and incorporated as appropriate, but new or more advanced concepts may be required because of the wide range of unknowns and uncertainties involved.

TABLE 12. Prioritization Factors for Improving and Evaluating Sensor Fusion Techniques

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	high	moderate	moderate	high

Sensor fusion has been targeted as a high priority for the Department of Defense (DoD) and the Defense Threat Reduction Agency. Research plans for various single or multiple approaches have been developed, and implementation of them has begun across a wide range of atmospheric scales. Candidate approaches can be evaluated in a relatively short time frame. Since the success of this R&D effort depends on the ATD modeling capabilities, sensor fusion capabilities should improve as ATD models improve. However, uncertainty due to the nonlinearity of ATD may limit capability improvements. Opportunities exist to understand this uncertainty and feed that understanding back into the ATD model improvement and sensor fusion capabilities. These synergies and dependencies across individual R&D needs are typical of what the JAG found in many areas for which quantifying ATD modeling uncertainties is a major objective.

4.4.2 Data Representation and Data Assimilation

Diverse sources of meteorological and agent-concentration data exist. The data are often taken at, or are representative of, varied time and space scales. Putting these data together in a coherent, physically realistic manner is the data assimilation process. As mentioned earlier, it may occur in one to four dimensions and cover a variety of scales. A key factor in data assimilation is how well the data fit the assumptions of the model: defined in section 3.2.3 as data representation. Prognostic models, for example, require initial conditions to describe the current state of the atmosphere in the volume being modeled and boundary conditions for that volume of atmosphere; namely, its inflow, outflow, and lower and upper sources or sinks of mass, energy, and momentum. The process of model initialization begins at global scales where radiosonde, satellite, aircraft, and other observational data sources are assimilated with previous forecasts for the new initial time to provide a comprehensive, updated, large-scale description of the atmosphere. As higher resolution grids are nested into subsets of the next larger scale, the initial and boundary conditions for the smaller grid must be provided to account for processes not represented at the larger scale.

At the larger scales, data assimilation approaches are undergoing intensive research and development. In the past two decades, physics-based assimilation of observations has replaced interpolation approaches. For example, in *nudging* techniques, a tendency term is added to the differential equation for each explicit variable in a full-physics mesoscale model such as MM5. The tendency term is proportional to the difference between the model predictions and observations; the observations are weighted in space and time, depending on the data type. Another approach comes from Sakaki's use of variational calculus, which opened the way to weighting observed data and constraints imposed by the equations of motion and conservation within the forecast model to produce three-dimensional initializations. This approach, called 3-D VAR, has become a de facto standard for major forecast centers. Attempts to incorporate temporal variability into the basic 3-D VAR approach, called 4-D VAR, lead to a complex set of equations that can only be solved approximately. Producing a 4-D VAR analysis is almost as time-consuming as the 72-hour forecast. Furthermore, it is closely coupled to the forecast model and its parameterizations.

More recently, ensemble techniques are being explored to improve the data representation fit of initial conditions. As noted previously, ensembles can be composed of several runs using the same model with different initial conditions, several runs using different models, or a set of runs with the same model but using different parameterizations. Another approach being explored for developing an ensemble is to use principles of Kalman filtering. These techniques appear to be independent of the forecast model. Generally speaking, most of these techniques have not been well tested with the finely nested grids needed for ATD predictions.

As grids are nested, typically in a ratio of 3:1, the meteorological data fields are usually interpolated to the finer grids without a physical constraint. Outside the ABL, this approach is reasonable. Inside the ABL, however, smaller scale processes that affect the turbulent state in the ABL are not represented. The usual assumption is that the finer-scale lower boundary conditions will develop (in the model state) after an initial adjustment period. Lateral boundary conditions are likewise interpolated and, at the inflow, driven by the larger scale forecast fields. Unfortunately, because there is a severe lack of observational data at nesting scales, these assumptions are not tested or directly verifiable. The overall forecast quality is a surrogate indicator of the validity of the assumptions. In traditional measures of performance, which use indicators such as bias and root-mean-square error, there appears to be a practical limit to improving forecasts (i.e., reducing forecast errors) by reducing grid size (Mass et al. 2002). This limitation can largely be explained by the fact that, as higher resolution is added to the forecast, the traditional forecast statistics become increasingly affected by slight errors in the location or timing of the mesoscale features. Research is needed to develop measures of performance for high-resolution mesoscale model forecasts that better quantify their value for use with ATD models.

Incorporating additional data on local winds and turbulence appears to have a positive effect on model performance at high resolution. As mentioned, the top-down approach of nesting appears to meet a practical limit near grid sizes of 5 km; however, the Army test ranges have had good results with operational modeling systems that use two-way grid interactions and an innermost domain resolution of about 1 km (Warner et al. 2004). For a wide range of weather conditions, models, and observational conditions, forecast models at that scale have large errors in wind-direction predictions—root mean square errors of about 40 to 60 degrees. Thus, inclusion of local observations in diagnostic and prognostic models of the wind field seems to be a reasonable and perhaps necessary approach to meeting user requirements for greater accuracy and useful information about predictive uncertainties.

Two other aspects of ATD modeling system performance place additional challenges on data assimilation R&D to improve the prediction of concentrations for the end users (emergency managers, operations officers, and researchers). First, for initial and intermediate response to hazard releases, data QA/QC for model fit (data representation) and data assimilation need to be automated (i.e., handled by software-embedded algorithms), without requiring expert “tweaking” by the model user. Second, the remote measurement technologies discussed in section 4.3.2 provide input data that require new capabilities on the part of the ATD model code to assimilate the data. Some software-embedded algorithms for data assimilation exist. However, as noted in section 3.2.3, assimilation of observations beyond t_0 of a prognostic model is often restricted by the constraints necessary to perform the iterated computations. If the data to be assimilated diverge too far from the model’s predicted values for that space-time cell, the data are rejected.

One of the emerging remote-sensing technologies, Doppler lidar, offers promise for providing high-resolution local wind fields in a variety of conditions. Following developments in deriving wind fields from Doppler radar data, Lin, Chai, and Sun (2001) used 4-D VAR with Doppler lidar data to construct three-dimensional wind fields characteristic of the convective boundary layer. As noted above, pure 4-D VAR is computationally time-consuming. Warner et al. (2002) used less restrictive constraints to permit a rapid analysis of 3-D wind fields obtained from scanning Doppler lidars.

Coupled with backscattered energy from airborne aerosols, almost real-time estimates of aerosol plume position, and short-term estimates of future paths (nowcasts) are feasible. Newsom and Banta (2004) and Calhoun et al. (2004) have suggested other approaches to assimilate lidar data at high resolution. Parallel development of lidar technology and data processing techniques should help advance the knowledge of smaller scale motions in real boundary layers.

R&D Need: Improve and evaluate techniques for data QA/QC for model fit and data assimilation for both initial and boundary conditions.

Data assimilation issues are closely tied to the scales of motion of interest, the availability of data representing those scales, and the techniques (models) used to link the data to the current state of the atmosphere at that scale. At present, data assimilation practices using variational or ensemble techniques exist for mesoscale operational models. These models are nested in global models but use finer resolution terrain conditions from surface, satellite, and/or aircraft regional observations. As finer scales are needed, assimilation approaches must adapt as surface and near-surface data become more important—an issue closely linked to improving characterization of surface-boundary conditions as discussed in section 4.2.2. At finer scales, assimilation becomes more temporally sensitive (perishable) and acceptant of observations appropriate to the model scale. The assimilation must allow representation of finer-scale dynamical processes (a need closely linked to bridging the mesoscale to microscale/urban scale gap, as discussed in section 4.2.2). It must be able to accept data coming from emerging measurement technologies (closely linked to improving boundary-layer atmospheric measurement capabilities, as discussed in section 4.2.3. These improvements are particularly important for recognizing and incorporating into the model run the three-dimensional structure of the daytime and night-time boundary layers.

Data QA/QC issues increase as remotely sensed and higher density in situ data are incorporated into the analyses. As much as possible, these issues should be addressed by onboard processing at the sensor, but errors due to data transmission, omissions, and losses must be identified before the data are used. Because volumetric remote sensing provides large sets of data to control and check, tests and filters for rapid and automated data QA/QC must be developed. Automated capabilities are also needed to ensure data representation by assessing the applicability of the data for the intended use. Including error bounds with observation data is an essential step toward understanding and quantifying the sources of uncertainty in model predictions that stem from factors outside the model itself.

TABLE 13. Prioritization Factors for Improving and Evaluating techniques for Data QA/QC for Model Fit and Data Assimilation, for Both Initial and Boundary Conditions

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
immediate	average	moderate	moderate	high

Substantial development has occurred in data QA/QC and data assimilation at the macroscale and the larger mesoscale. Applications of 4D VAR and ensemble techniques for meso–β and smaller scales can be initiated without much difficulty for testing these techniques where appropriate measurements exist. Full implementation of assimilation techniques will become beneficial when data become more plentiful and regular in test beds or large networks (i.e., when instrumentation is developed and in operation). As model scales become smaller and approach the urban scale, data assessment issues will become more location-specific, adding challenges to automation of the process and requiring faster execution times. Results of this R&D will link with the design and implementation of urban regional monitoring networks.

4.4.3 Model Performance Evaluation Issues

Model performance evaluation basically comes down to comparing a model's predictions of concentration with the concentrations observed from field measurements. One can view the observed concentrations as a summation of three values: the ensemble average for the conditions present, the effects of measurement uncertainty, and the effects of unresolved processes (stochastic fluctuations). The modeled concentrations can be viewed as a summation of three values: the ensemble average for the conditions present, the effects of uncertainty in specifying the model inputs, and the effects of errors in model formulation (which may vary as conditions vary).

The concept of natural variability acknowledges that the details of the stochastic concentration field resulting from transport and diffusion are difficult to predict. In this context, the difference between the ensemble average and any one observed concentration value (realization) is ascribed to natural variability. The ensemble is the ideal infinite population of all possible realizations meeting the (fixed) characteristics associated with the ensemble. In practice, one will only have a small sample from this ensemble.

Measurement uncertainty in concentration values in most tracer experiments may be a small fraction of the measurement threshold. When this is true, the contribution of the measurement uncertainty to empirical determinations of the magnitude of natural variability can usually be deemed negligible.

One method for performing an evaluation of modeling skill is to average separately the observations and the modeling results over a series of non-overlapping limited ranges of fixed conditions, which are called "regimes." Averaging the observations provides an empirical estimate of what most of the current models are attempting to simulate; namely, the ensemble average. A comparison of the respective observed and modeled averages over a series of regimes provides an empirical estimate of the combined error associated with input uncertainty and formulation errors.

This method for evaluating model skill is not perfect. Some models provide estimates of the average concentration for a volume of air (grid averages), whereas the observations represent what is seen for some point in the volume of air. The variance in observed concentration values due to natural variability can be on the order of the magnitude of the regime averages. Hence, small sample sizes in the groups will lead to large uncertainties in the estimates of the ensemble averages. The variance in modeled concentration values due to input uncertainty can be quite large; small sample sizes in the groups will therefore lead to large uncertainties in the estimates of the deterministic error in each group. Finally, grouping data together for analysis requires large data sets, of which there are few.

R&D Need. Develop physics-based evaluation metrics that recognize the fundamentally different sources for variations in observed and model-predicted values of downwind hazard concentration.

The most important concept expressed in the discussion above of modeling performance evaluation is that the observations and the modeling results come from different statistical populations whose means are (for an unbiased model) the same. The variance seen in the modeled values results from differences between estimates of ensemble averages and differences resulting from modeling errors. The variance in the observations results from differences in ensemble averages, differences arising from sampling uncertainties, and an additional variance, which is not represented in deterministic modeling, caused by stochastic variations between individual realizations. Because of these differences in the populations for which variances are being estimated, a thorough reassessment is needed of how transport and diffusion models are evaluated. The currently accepted model evaluation methods directly compare the observed and modeled concentration values (in contrast to comparison of regime averages), an approach that assumes the observations and the model estimates have the same sources of variance. As explained above, this assumption is erroneous. Viewed in this context, comparisons of observed and modeled frequency distributions of concentration values for transport and diffusion models are questionable, unless the models are attempting to estimate not only the variations to be expected in the ensemble average as conditions vary but also the effects of unresolved stochastic fluctuations. Thus, asking whether a deterministic model can match observed extreme values amounts to requiring the model to succeed at a task it is fundamentally incapable of doing, except by compensating for input uncertainties and formulation errors. Until now, model evaluations have focused on evaluating a model for how it is used rather than on the basis of what the physics in the model is capable of estimating. For example, models are now often evaluated as a characterization of extreme values—a task for which few, if any, models incorporate the necessary physics.

Thus, the focus of model evaluation methods should be on assessing how well a model predicts those features of the concentration distribution (mean, variance, distribution) for which that model incorporates appropriate physics. While we cannot simulate exactly what is observed in time and space, we might (with suitable research) predict the average characteristics of the concentration distribution seen at each point (e.g., the mean, variance, and distribution). Of course, we only observe individual realizations, but if we properly predict the characteristics, the observed individual realizations will be within the predicted distribution of possible outcomes. If this approach to model evaluation is pursued, the evaluation methods can adapt to assess model performance as new model capabilities (e.g., probabilistic modeling) are developed.

TABLE 14. Prioritization Factors for Developing Physics-based Model Evaluation Methods

Time Sensitivity	Short-Term Gain	Overall LOE	Lead Time	Ultimate Gain Potential
near term	high	low	average	exceptional

Physics-based model evaluation methods must consider the internal, external, and stochastic components of model uncertainty. Models are evaluated unevenly, with different criteria applied by different developers. A reference standard or consensus-based methodology developed by an independent standards setting organization provides a standard by which developers and evaluators can uniformly evaluate modeling systems. This solution can be implemented rapidly by commissioning a standard-setting organization to develop and maintain (update) the standard. The sustaining activity by the organization will ensure the standard is maintained over time as experience is gained and innovation produces improvements.

4.5 Summary of R&D Needs

Table 15 is a compilation of the prioritization factors assigned to R&D needs in tables 3 through 14. Although this summary table brings all the R&D needs together, the assignments of prioritization factors need to be interpreted through the explanations given in the paragraphs following each of the component tables.

TABLE 15. Summary Table of R&D Needs with Prioritization Factors

R&D Need	Time Sensitivity	Short-Term Gain	Overall Level of Effort	Lead Time	Ultimate Gain Potential
Bridge the modeling gap	near term	average	moderate	average	exceptional
Characterization of surface conditions & input data sets	near term	average	high	average	exceptional
Test and refine physical basis for sub-grid-scale parameterizations	longer term	average	moderate	average	exceptional
Characterize dispersion in complex environments	immediate	average	high	average	high
Improve ensemble construction and interpretation	immediate	minimal	high	short	exceptional
Techniques to better estimate wet and dry deposition	near term	average	moderate	average	high
Physical and high-resolution computational models	near term	average	moderate	average	exceptional
Improve tracer materials and measurement technology	immediate	high	moderate	short	exceptional
Improve boundary-layer measurement technology	immediate	high	high	short	exceptional
Improve and evaluate sensor fusion techniques	immediate	high	moderate	moderate	high
Data QA/QC for model fit and data assimilation	immediate	average	moderate	moderate	high
Develop physics-based model evaluation methods	near term	high	low	average	exceptional